

Common Weakness Scoring System (CWSS)

Steve Christey

September 27, 2010

cwss@mitre.org

<http://cwe.mitre.org/cwss>



The Problem

- **In the process of discovering new vulnerabilities, automated and human analysis will find weaknesses**
 - Everyone scores weaknesses differently
- **Not all reported weaknesses necessarily indicate a vulnerability**
- **Hundreds or thousands of weaknesses could be reported for a single software package**
- **Weaknesses can be treated as an entire class of problem to eradicate, independent of any specific bug in a specific software product**
- **Support for a variety of contexts and threat environments**

Beginnings

- **CWSS Kickoff Meeting – Oct 24, 2008**
- **Briefing to SwA Working Groups – July 2010**
- **Start with CVSS**
 - Try to address some of CVSS' limitations
 - Examine other metrics
- **Environment / Context is critical**
 - Business/mission priorities, how SW is deployed, ...
- **Ideally supports tools and humans**
- **Must be stable/predictable even when there is limited information**

CWSS Kickoff Meeting

- **October 24, 2008**
 - ~12 organizations represented
 - Primarily tool vendors and consultants
- **Main topics**
 - What weakness-scoring approaches are currently used?
 - How much should CWSS borrow from previous efforts like CVSS?
 - How closely should CWSS and CWE be affiliated?
 - Who are the stakeholders?
 - What are some core elements to capture in CWSS?
- **Main conclusions**
 - CWSS is independent of CVSS
 - Complete information is rarely available
 - Multiple stakeholders are involved
 - Distinct usage models
 - Line between “Weakness” and “Vulnerability” is fuzzy



Related Scoring Efforts (from 2008 kickoff)

- **Veracode – Confidentiality, Integrity, Availability analysis of CWE weaknesses**
 - Independence from analysis method
- **Cigital – feasibility study of CVSS**
 - Some attributes like “Target Distribution” didn’t fit well
 - Added more granularity to some attributes
 - Recommended polynomial scoring
 - Important to model the distinction between likelihood and impact

Related Scoring Efforts (Continued)

- **Cenzic – HARM (Hailstorm Application Risk Metric)**
 - Quantitative score focused on black box analysis of web applications
 - Goal: scalable focus on remediation
 - 4 impact areas: browser, session, web app, server
 - Benefit: “easily consumable”
- **CERT/SEI – scoring of C Secure Coding Rules**
 - FMECA (ISO standard) metric: Severity, Likelihood (of leading to vuln), Remediation Cost

2010 SANS/CWE Top 25

- **Real-world, raw data is still very difficult to find**
- **Prioritized items based on “Prevalence” and “Importance” (4 values each)**
- **25 participating organizations evaluated 41 nominee CWE entries**
 - Developers, researchers, educators
- **Focus profiles allowed alternate ranking**
 - E.g. educational emphasis, importance to software consumers

2010 OWASP Top Ten - Factors

- **Ease of Exploit**
- **Prevalence**
- **Detectability**
- **Technical Impact**

2009 CWE Top 25 - Informal Criteria

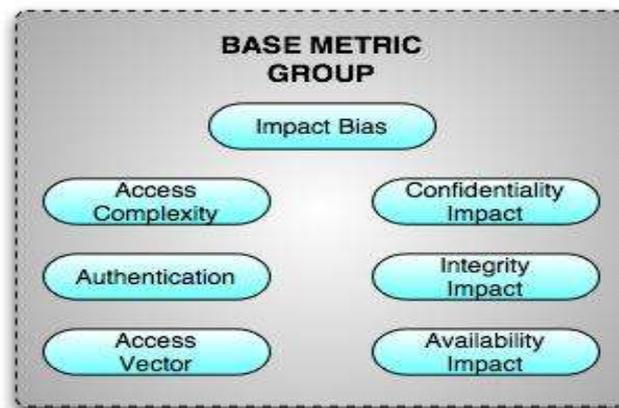
- **Attack frequency**
- **Consequences**
- **Weakness prevalence**
- **Ease of detection**
- **Language/platform independence when possible**
- **Others considered: remediation cost, amount of public knowledge, likelihood of future increase**

Common Vulnerability Scoring System

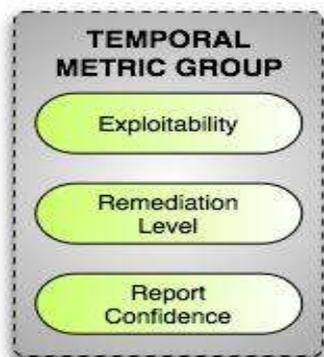
- NVD provides CVSS scores for all CVE identifiers
- Payment Card Industry (PCI) using CVSS to determine compliance
- Formal validation program (~10)

www.first.org/cvss/

Core aspects of the problem



Reflects your own organization's priorities



Changes over time



Some CVSS Strengths

- **Relative Simplicity**
- **Repeatability (within some epsilon)**
- **Efficient representation (vector)**
- **Widely adopted**
- **“Good enough” for non-expert admins**

Some CVSS Limitations

- **Focuses on impact to system**
 - The “Oracle” problem
- **Requires good documentation**
- **Not granular enough for expert consumers**
 - E.g. confidentiality/integrity/availability
- **Doesn't handle insufficient information well**
 - The “Missing Oracle” problem
- **Temporal/Environmental aspects not well-tested**

Some Potential Stakeholders for CWSS

- **Software developers/programmers**
 - “We’ll concentrate on what we can afford to fix, or what our worst problems are”
- **Software project managers**
- **SW acquirers**
 - “The purchased software shall not have any outstanding weaknesses greater than CWSS score 7.0, as determined by methods X and Y.”
 - Large enterprise organizations
- **Code analysis vendors – tools and services**
- **Vulnerability researchers**
- **Secure development advocates**
- **CIO’s and CSO’s**
- **System administrators**
- **Application users**



Developer Use Case: Example

- ***I ran a tool that gave me hundreds of results. Which issues should I address, based on:***
 - My company's goals
 - Compliance requirements
 - Customer expectations
 - Best current practices
 - Amount of time before next release
 - My company's willingness to accept the risk of not fixing

Support for Multiple Scoring Modes

- **General (Prevention): score weaknesses based on their general occurrence in software**
 - In general, how bad are buffer overflows versus memory leaks?
 - Won't always be correct for a specific instance
 - “Top 25,” and other lists
 - But even “buffer overflow” risk varies widely, e.g. overflow protection
- **Specific (Reaction): score a weakness based on its occurrence within a specific software package**
 - Won't always have complete information

Scoring Modes

- **General** – “in general, how bad are buffer overflows versus memory leaks?”
- **Specific** – “how to score weakness X in line 1234 of vuln.c”?
- **Use maximum possible score, adjust downward given environmental/contextual information**
 - “this is a config file only readable by admin”
 - “this input is externally validated using Struts”
- **For a software package, how to combine all reported weaknesses to get an overall score?**

Vignettes

- **Define a particular environment and its priorities**
 - Essential Resources / Capabilities
 - Confidentiality/Integrity/Availability importance
- **Define priorities for each general “Technical Impact”**
 - E.g., read application data, execute code
 - These are in each CWE entry
- **Initial focus on CWE/SANS Top 25**

Vignettes – Initial Focus

Name	Description	Resources / Capabilities	Notes
Web-based Retail Provider	E-commerce provider of retail goods or services through a web-based interface.	data-centric - Database containing PII and inventory. Based on web server.	Confidentiality essential from a financial PII perspective, identity PII usually less important. PCI compliance a factor.
Financial Trading/Transactional	Financial trading system supporting high-volume, high-speed transactions	n-tier distributed, using J2EE and supporting frameworks. Core business logic is J2EE. Transactional engine between multiple data feeds; no core "front end."	High on integrity - transactions can't be modified. Availability also very high - if system goes down, financial trading can stop. Critical transactions not processed.
SCADA-based flow control system for chemical plant	Controllers for physical systems in "critical infrastructure" e.g. power plants, hydro, chemical production, etc. Compromise could result in ecological disaster, explosions, poison.	underlying technology - heavy C usage. Systems developed in pre-Internet era with management consols interfacing them.	Confidentiality matters little if at all. Availability - for a control system, high. For a monitoring system, less so. Integrity can be very high - does this actually control/modify physical systems.
Human Resources	Product for managing human resource data, including recruitment, compensation, benefits, performance evaluations, and training.	data-centric - SOA-based web service; heavy DB applications. highly centralized PII related to identity and "personal history" information.	Protection of private data is critical to minimize exposure to lawsuits, adverse impacts on morale.
Social networking	Product for enabling a large community of people to chat, exchange messages or pictures, and share interests, e.g. Facebook, MySpace, Twitter, LinkedIn.	Service-oriented product. Heavy Web 2.0. Free-form content, high connectivity between users, privacy considerations (decentralized), lots of messaging.	Privacy concerns regarded as less important than integrity of the data; uptime an important concern.

Vignette-Oriented Scoring

- Calculate “General” scores and “Vignette-Oriented” scores
- Focus on CWE/SANS Top 25
- Borrow Prevalence data from Top 25
 - Normalize between 1 and 10
- Importance may vary per Vignette

CWSS 0.1 Scoring

Prevalence * Importance

CWE 1.10 – Technical Impact

- **Modify memory**
- **Read memory**
- **Modify files or directories**
- **Read files or directories**
- **Modify application data**
- **Read application data**
- **DoS: crash / exit / restart**
- **DoS: amplification**
- **DoS: instability**
- **DoS: resource consumption (CPU)**
- **DoS: resource consumption (memory)**
- **DoS: resource consumption (other)**
- **Execute unauthorized code or commands**
- **Gain privileges / assume identity**
- **Bypass protection mechanism**
- **Hide activities**

Each vignette requires analysis of these technical impacts; *not* individual CWE entries.

Vignette-Oriented Scoring (2)

- **Calculate Vignette-Specific Prevalence**
 - For CWSS 0.1: just re-use Top 25 voting data
- **Calculate Vignette-Specific Importance**
 - For each vignette:
 - Evaluate each CWE Technical Impact and assign factors
 - Examine each CWE entry, calculate subscore for each listed Technical Impact in that entry
 - Choose max of all subscores
 - Normalize to between 1 and 10? 1 and 100?

Technical Impact Analysis: Example

Technical Impact	Base Subscore (Max)	Base Subscore (Ave)	Financial Trading	Human Resources	Retail WWW	SCADA / Chemical	Social Network
Modify files or directories	8	7.60	subscore: 8 delete transactions ; inject fraudulent transactions ; remove transaction history.	subscore: 8 modification of salary, identity PII (e.g. social security #), potentially medical information (e.g. long-term health issues), disciplinary history	subscore: 8 modify customer orders; deface web pages; install malware through web pages; modify system configuration	subscore: 6 present incorrect information to operators, leading to wrong decisions; "remove" controllers from network, preventing them from being monitored at all.	subscore: 8 Falsification of user profiles, contact information, status updates; deletion of account data, message history, contact information, etc.

Technical Impact Analysis: Example 2

Technical Impact	Base Subscore (Max)	Base Subscore (Ave)	Financial Trading	Human Resources	Retail WWW	SCADA / Chemical	Social Network
Read application data	8	5.00	subscore: 8 enabling insider trading; breach confidentiality of transactions (knowing that someone's doing something)	subscore: 8 read salary, employment history, identity PII (e.g. social security #), potentially medical information (e.g. long-term health issues), disciplinary history	subscore: 2 exposure of customer CC numbers (see TJX). Resulting impacts - financial restitution, legal liability, compliance/regulatory concerns, reputation/brand damage.	subscore: 1 reconnaissance for a different type of attack	subscore: 6 basic PII (phone, email, address, location), affiliations with other people, reading private communications, ...

Individual Vignette Breakdown: Web-based Retail Provider

Technical Impact	Subscore	Description
Hide activities	3	Inability to identify source of attack; cannot perform legal prosecution
DoS: resource consumption (CPU)	2	Customers experience delays in difficulty reaching site; financial loss due to delays in order placement
Modify application data	7	modify customer contact data (physical/electronic addresses); install malware; corrupt / modify order histories
Read application data	5	exposure of customer CC numbers (see TJX). Resulting impacts - financial restitution, legal liability, compliance/regulatory concerns, reputation/brand damage.

Technical Impacts for Individual CWE Entries - Example

- **CWE-89: Improper Neutralization of Special Elements used in an SQL Command ('SQL Injection')**
 - Read application data
 - Bypass protection mechanism
 - Modify application data
- **CWE-120: Buffer Copy without Checking Size of Input ('Classic Buffer Overflow')**
 - Execute unauthorized code or commands
 - DoS: crash / exit / restart



Prevalence Assessments per CWE

CWE	Name	Prevalence (1-10)
CWE-79	XSS	9.46
CWE-89	SQL injection	7.43
CWE-120	Classic overflow	6.04
CWE-352	Cross-site Request Forgery	7.75
CWE-285	Insufficient Authorization	6.04
...

- Borrow data from 2010 CWE Top 25 votes
- Normalize 1-4 scores to 1-10 range

CWSS 0.1 Scoring Examples

CWE	Name	Prevalence	Retail WWW	Financial Trading	SCADA / Chemical	Human Resources	Social Network
CWE-79	XSS	9.46	94.6 (imp=10)	94.6 (imp=10)	94.6 (imp=10)	94.6 (imp=10)	94.6 (imp=10)
CWE-89	SQL injection	7.43	59.4 (imp=8)	59.4 (imp=8)	59.4 (imp=8)	59.4 (imp=8)	59.4 (imp=8)
CWE-352	CSRF	7.75	69.8 (imp=9)	62.0 (imp=8)	69.8 (imp=9)	62.0 (imp=8)	62.0 (imp=8)
CWE-209	Information Exposure Through an Error Message	7.11	14.2 (imp=2)	56.9 (imp=8)	7.1 (imp=1)	56.9 (imp=8)	42.7 (imp=6)

Calculating the General Score

- **For each CWE**
 - For each Technical Impact, get the max Importance of all vignettes
 - Alternate: choose the average
 - Use this as the base Importance
 - Use Top 25 data for base Prevalence
 - $\text{Prevalence} * \text{Importance}$ would range from 1 to 100
- **Since there is a limited set of Technical Impacts, this is much more efficient than manually evaluating importance of each individual CWE**

Considerations for General Scoring

- **Factors for general score will change regularly**
 - Prevalence may change
 - Vignettes may change
 - “versioning” for CWSS when factors change
 - Is this manageable when sharing CWSS scores?
- **If maximum impact is chosen instead of the average, then scores will trend upwards as more vignettes are added**
- **Scores could be adjusted downward based on environmental considerations**
 - Still need to model these

Possibilities for Specific Scoring

■ Inherent Risk

- Known-exploitable vs. unlikely? (e.g. buffer overflows)

*Known to
tools*

■ Inherent Attack Complexity

- Level of authentication required
- Local/remote
- Amount of victim interaction required

*Known to
developers*

■ Execution Environment

- Specific to some or all configurations? Default?

■ Threat Environment

Script kiddies or nation-state? Financial or physical?

*Known to
customers*

■ Business/Mission Priority

- “Top 25” list conformance, customer needs, software vendor training/education

Future Activities

- <http://cwe.mitre.org/cwss>
- Upcoming white paper
- Build community
 - Software vendors
 - Automated code analysis vendors
 - Current CVSS community
 - Academic community
 - Researchers/consultants
- Solicit feedback for Vignettes
- Clarify CWE “Technical Impacts” model
- Use of CWSS 0.1 in Pocket Guide for addressing the Top 25

cwss@mitre.org

BACKUP SLIDES

Example: Prevention Mode

- **General Exploitability**
 - Level of attacker knowledge of the weakness
 - Availability of tools that scan for the weakness
 - Availability of tools that attack the weakness
 - Frequency of exploitation in the wild
 - Likelihood of future popularity / attacks (e.g. CSRF)

Example: Reactive Mode

- **Specific Exploitability**
 - Proven vs. likely (e.g. buffer overflows)
 - Reachability
 - Maybe it's not currently reachable
 - dead code
 - all code paths are currently safe
 - Discoverability
 - Indirect protection mechanisms

Inherent Risk

- **Consequences**
 - Confidentiality, Integrity, Availability, Non-Repudiation, ...
 - Code execution, data corruption, information leak, ...
- **Relative to:**
 - Application's users
 - Application
 - Host system
 - Host network
- **Challenge: ~16 combinations for each weakness**



Inherent Attack Complexity

- **Access Requirements**

- Amount of interaction
- Privileges required

cwe.mitre.org/cwss/

- **Authentication**

- None, one, multiple – consideration of strength?
- What if the weakness is about authentication?

- **Access Vector**

- Local, Adjacent Network, Network, Physical, Virtual

- **Amount of interaction**

- None, typical user behavior, significant interaction
- E.g. “clicking on a link” in a web page is typical

Execution Environment

- **Execution context**
 - Is it only accessible to an admin?
 - Protection mechanisms e.g. overflow protection?
- **Application's role with respect to:**
 - The organization
 - The physical host
 - The network
 - The surrounding software system



Business/Mission Component

- **Organizational prioritization**
 - Management says, “no more buffer overflows”
 - Customer requires assurance level X
 - Weakness Z “could put us out of business”
 - Concentrate on components A and B
 - Risk acceptance level
- **Remediation Cost**
 - Most stakeholders don’t care (e.g. customers)
 - Length of time to fix (e.g. design vs. implementation)
- **Collateral Damage Potential (CDP)**
 - Financial
 - Physical
 - Personal Information
 - Software users
 - System users



Other Possible Factors

- **Report confidence**
 - Likelihood of true/false positive
 - Could be from a tool or a human
- **Target distribution**
 - Number of affected customers
 - Number of affected versions

Summary of CWSS Requirements

- **Incomplete information is the norm**
- **Environmental/business/mission must be considered**
- **Support for multiple scoring modes**
 - General vs. specific
- **Stakeholder analysis needed**
- **Metric complexity is a risk**